

1974

# The experimenter as an audience: the effects of varied experimenter "presence" upon performance at a simple laboratory task.

Grant M. Ingle

*University of Massachusetts Amherst*

Follow this and additional works at: <https://scholarworks.umass.edu/theses>

---

Ingle, Grant M., "The experimenter as an audience: the effects of varied experimenter "presence" upon performance at a simple laboratory task." (1974). *Masters Theses 1911 - February 2014*. 1631.

Retrieved from <https://scholarworks.umass.edu/theses/1631>

This thesis is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses 1911 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

UMASS/AMHERST



312066013803011

THE EXPERIMENTER AS AN AUDIENCE: THE  
EFFECTS OF VARIED EXPERIMENTER "PRESENCE"  
UPON PERFORMANCE AT A SIMPLE LABORATORY TASK

A Thesis Presented

By

Grant M. Ingle

Submitted to the Graduate School of the  
University of Massachusetts in partial  
fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 1974

Department of Psychology



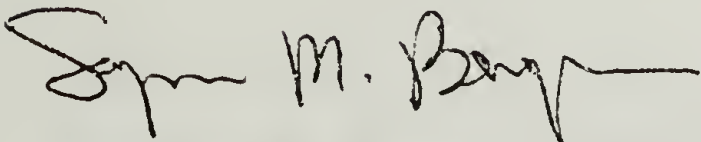
THE EXPERIMENTER AS AN AUDIENCE: THE EFFECTS  
OF VARIED EXPERIMENTER "PRESENCE" UPON  
PERFORMANCE AT A SIMPLE LABORATORY TASK

A Thesis

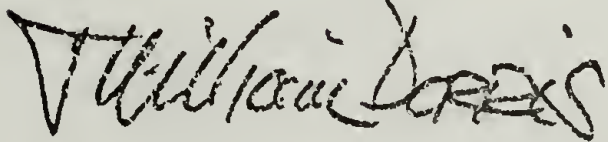
By

GRANT M. INGLE

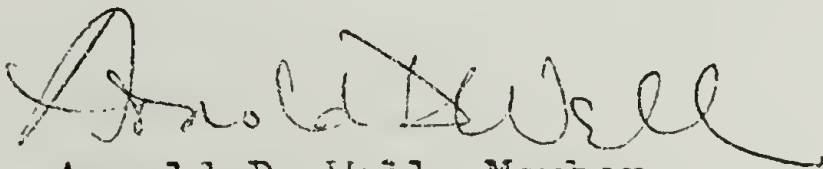
Approved as to style and content by:



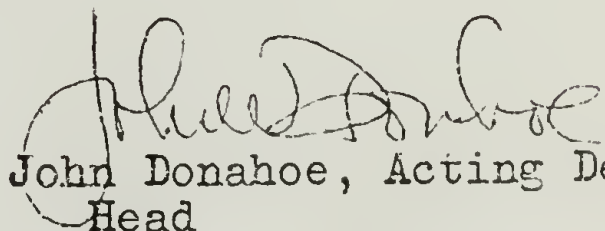
Seymour M. Berger, Chairman of Committee



J. William Dorris, Member



Arnold D. Weil, Member



John Donahoe, Acting Department  
Head  
Department of Psychology

August 1974

## ACKNOWLEDGEMENTS

At this time I would like to thank Dr. Seymour M. Berger for his patient assistance in the preparation of this thesis. His excellent criticisms and thoughtful guidance have been greatly appreciated. In addition, thanks are extended to Drs. Arnold D. Well and J. William Dorris, both members of my committee, for their comments and encouragement.

I would also like to thank Dr. Kenneth T. Hill, who was primarily responsible for the germination of the ideas contained herein, and also Cindy Doran, for her persistent typing.

## TABLE OF CONTENTS

Introduction. . . . .	page 1
Method. . . . .	page 17
Results. . . . .	page 21
Discussion. . . . .	page 27
Summary. . . . .	page 39
Footnotes. . . . .	page 41
Bibliography. . . . .	page 42

## LIST OF TABLES

Table 1 - Means and Variances for the Number of Vowels  
Cancelled During A Five Minute Interval. . . page 22

LIST OF FIGURES

- Figure 1 - Frequency polygon for distribution of scores in the visual monitoring condition versus aural monitoring conditions (with adjusted no monitoring scores). . . . .page 26



## Introduction

In 1919, Floyd Allport wrote:

When social psychologists focus their attention upon the behavior of the individual under direct and incidental stimulation from the behavior of others, then the most vital problems of the social order will find their solution (p. 305).

Allport's prescriptions have had a considerable impact upon much of the research which was to follow him in social psychology, especially in the area of social facilitation. Despite Allport's emphasis upon the "stimulation from the behavior of others," however, it is quite clear that the experimenter in laboratory research was not considered to be one of these "others." At the time that Allport wrote, this apparent oversight was actually quite consistent with the scientific model upon which emerging twentieth century psychology was built: nineteenth century physics. The most crucial assumption of nineteenth century sciences was, of course, that the observer (and thus the process of observation) is independent of what is being observed, or, "in a literal sense [that] the researcher is independent of his experimental situation" (Giorgi, 1970, p. 166). Given this initial assumption, then, it should be evident that early social psychologists acted quite logically in attaching little or no importance to the presence of the experimenter in the psychological laboratory. As a consequence of this development, however, contemporary psychology has become an unknowing heir to an "experimental situation" which often presumes the "presence" of the experimenter. This presence

is most often an actual physical presence which greets the subject, introduces the experiment, administers instructions, directs the task, collects each individual's data, and disburses money, credits and debriefings. Often this situation also involves an immediate monitoring of the subject's ongoing behavior, some aspect of which the experimenter finds important. If this monitoring is not accomplished by the physical presence of the investigator, then it frequently takes some electronic or displaced form, as in the case of intercoms, videotape, or one-way mirrors, etc. In any event, few psychologists have ever concerned themselves with this "presence" and its possible implications, theoretical or practical. Historically, however, at least some of the difficulties associated with experimenter presence were acknowledged periodically within the developing psychological literature.

The first prominent statement of the issue of experimenter presence appeared in the second volume of the Handbook of Social Psychology in an article by J. F. Dashiell (1935) who reviewed a number of studies which purported to examine the effects of spectators upon individual performance. These studies incorporated a general design which compared subject performance when knowingly observed by an audience with performance when the subject was alone. Few of the studies demonstrated statistically significant audience (social facilita-

tion) effects, although definite positive trends were evident. Close examination of the procedures employed in these studies disclosed that the experimenter was present with the subject in both the "audience" and "alone" conditions. Dashiell succinctly pointed out, "it is not to be rashly assumed that he [the experimenter] may not influence the subject as much as the spectator does." He then proceeded to further examine three studies which included the "experimenter himself as a social object (p. 1104)." In a study of mechanical and social distractors, Pessin (1933) found that the presence of the experimenter interfered with the memorizing of nonsense syllables to the same degree as flashing lights combined with a loud buzzer. Secondly, Dashiell related a study by Ichheiser (1930) which demonstrated that under the observation of the experimenter, subjects exhibited a decrement in performance at a block-assembling task when compared to a situation in which the experimenter was absent from the test room. In reference to these studies Dashiell commented: "If, now, the presence of the examiner is of importance in a routine mechanical task, shall we not expect it to mount higher when the task is one involving associational functions, particularly if other personality processes intrude upon the association processes? (p. 1105)" Supporting this expectation, Dashiell then reported a study by Ekdahl (1929), which demonstrated longer reaction times for word associations in conditions in which the experimenter was present versus absent.



Unfortunately, few other psychologists have shared the concerns of Dashiell. As a consequence, the only reports of experimenter presence phenomena have been those which inadvertently resulted from independent research in widely differing areas of psychology. In the area of signal detection, Fraser (1953) first demonstrated that the presence of an experimenter significantly increased performance relative to conditions in which the subject performed alone. In a later, but related study of the vigilance performances of enlisted army personnel, Bergum and Lehr (1963) combined their data with Fraser's and concluded that the presence of an experimenter caused increases in performance equal in magnitude to those induced by the presence of an army officer. A number of social reinforcement studies with children demonstrated a very similar effect. When the task involved performance at a simple task, such as marble-dropping, the presence of a non-reactive experimenter increased performance rates relative to conditions in which the experimenter was absent (Meddock, Parsons and Hill, 1971; Leventhal and Fischer, 1970; Peterson and Whitehurst, 1970). And, as Milgram (1965) found, "obedience dropped sharply as the experimenter was physically removed from the laboratory." Given the inheritance of an experimental situation which often presumes the physical inclusion of the experimenter, it is not surprising that these findings arose incidentally to the major foci of these researches; this phenomenon was treated as an oddity of experi-

mental designs and tasks, rather than being interpreted as a broad methodological problem generally applicable to the experimental paradigm.

Ironically, nowhere has this disregard for experimenter presence been more apparent than in the study of human social facilitation effects, as evidenced by the large number of recent studies in which this consideration necessarily would be critical. In the more recent literature, for example, three studies which contain audience/alone manipulations, but which also fail to demonstrate an audience effect, all employ the same questionable procedure: the experimenter is physically present in both the "alone" and "audience" conditions (Shrauger, 1972; Criddle, 1971b; Criddle, 1971c). Perhaps the most blatant example of this oversight is afforded by the research reported by Shrauger (1972). In this study, the "audience" condition was created by introducing and seating two psychologists in the same room with the subject and the experimenter; the "no audience" condition consisted of only the subject and the experimenter. In discussing the failure of this manipulation to produce a significant effect, the author makes no mention of the possibility that his presence itself may have created an audience effect, thus nullifying any effect due to the addition of two more psychologists.

Even more disconcerting, particularly at the theoretical level, are similar studies which have reported significant effects, but at the same time have included the physical pres-



ence of the experimenter across all audience conditions (Martens, 1969; Wapner and Alper, 1952; Ganzer, 1968). The issue becomes especially confusing in studies like that of Cottrell, Rittle and Wack (1965), in which the experimenter was always present with the subject and both positive and null effects were found for different audience conditions. In this study, for example, "mere presence" (blindfolded peers) was not found to create a significant degree of social facilitation effects when compared with the "alone" condition, but, significant effects were found when an unblinded peer audience was present. In this situation, then, by careless definition of the subject "alone" condition, the social facilitation effects found in that "alone" condition may actually have been inflated by the observing experimenter.

A reinterpretation of other social facilitation research can provide a critical examination of the specific effects of experimenter presence, even though this research was not conceived with this end in mind. Initially, at least, this notion of experimenter presence coincides nicely with Zajonc's (1965) requirement for the occurrence of social facilitation effects. Specifically, Zajonc has suggested that the "mere presence" of others will facilitate a well-learned response, but would impair the acquisition of new or subordinate responses. More recent consideration of the problem, however, has seriously questioned Zajonc's requirement of "mere physical presence." Hanchy and Glass (1968) explicitly contend that the arousal

of evaluation apprehension is the necessary condition for the occurrence of social facilitation effects, while this position is implied by other researchers such as Cottrell (1968), who suggests that the "anticipation of positive or negative outcomes" is the necessary requirement. Examining the problem in terms of the more distal<sup>1</sup> determinants of the phenomenon, Criddle (1971a) has maintained that "some sort of monitoring" is the necessary condition, and thus has questioned the requirement of "presence" altogether. Regardless of the particular interpretation of social facilitation phenomena, one conclusion is inescapable: the experimenter's presence in the experiment must be examined as a special case of social facilitation effects; the experimenter is an audience who monitors the subject's behavior.

In this regard, it is important initially to define conceptually the meaning of "experimenter presence" and this will demand a somewhat detailed examination of two of the more recent social facilitation studies. Criddle (1971a), in testing Zajonc's "mere presence" requirement, used a paired associates task with either competitive (difficult) or non-competitive (simple) lists to assess the social facilitation effects arising from two different conditions. Subjects in the "audience" condition were led to believe that they were being observed through a one-way mirror, while in the "no audience" condition, the mirror was covered. After the instructions were presented to the subjects, the experimenter

left the room and the task was presented by means of a tape recorder; the subject's responses were recorded on a second recorder. Criddle demonstrated that performance at the competitive list was impaired by subject knowledge of observation, but failed to find evidence for the facilitation of performance at the non-competitive list (though it appears that his failure was probably due to a ceiling effect inherent in the task). Although the results were somewhat weak, Criddle's study suggests that subjects who are aware that their behavior is being monitored will demonstrate social facilitation effects, a conclusion which is compatible with the results of Henchy and Glass (1968). Although Criddle does not extend his finding to experimenter monitoring specifically, the implications are quite clear: social facilitation effects due to the presence of the experimenter may not require that experimenter's physical presence within the experimental setting. With regard to Criddle's study (1971a), for example, it should be emphasized that the experimenter was physically absent (after presenting instructions) in both the "audience" and "alone" conditions, and that "monitoring" by the experimenter was confined to the tape-recording of the subject's responses. In two other related studies in which the experimenter was present in all conditions (Criddle, 1971b; Criddle, 1971c), the same author failed to find significant audience effects between different "monitoring" conditions. In the first of these studies, subjects performed a pseudo-



recognition task in one of three varied conditions of observation: (1) peers physically present; (2) peers (supposedly) watching from behind a one-way mirror; and (3) no audience other than a tape recorder. Criddle had predicted results consistent with his "monitoring" approach (more monitoring = greater audience effects): physically present peers should produce the strongest effects, followed next by the one-way mirror audience, and finally by the tape recorder. Contrary to this hypothesis, Criddle reported that there were no significant differences between any of the conditions. This lack of results, then, suggests a "basement effect": different "monitoring" conditions are "washed out" when combined with the experimenter's physical presence, since his presence itself may be a stronger condition of observation than any of the experimental treatments. In other words, the subjects in this study were more concerned about the physical presence of the experimenter than they were about other ways in which their task performance might be monitored.

Given the perspective that the experimenter serves as an audience in the laboratory, it appears reasonable to suggest that experimenter monitoring can occur to differing degrees, depending upon the particular observational procedure employed in a given experiment. Actually, the use here of the term "experimenter monitoring" may be conceptually misleading, in that the term is derived from the experimenter's point of view and thus concentrates solely upon the experimenter's behavior

in the setting. Phrasing the problem with regard to the subject's perspective, on the other hand, appears to enhance an understanding of the situation. Given this orientation, it might be said that the observational procedure of a given experiment determines the degree to which the subject is aware of being monitored by the experimenter. Clearly, then, different degrees of experimenter monitoring are not limited solely to variations in the physical presence or absence of the experimenter. These variations constitute only one way in which the subject's awareness of being observed may be altered. The presence of video cameras, microphones and intercoms are also blatant indicators to the subject that "monitoring" is occurring. Their obtrusiveness, however, may be of a lesser magnitude than that of the experimenter's presence, in that the experimenter, by eye contact, movement, noise, etc., may continually remind the subject that his or her behavior is under immediate scrutiny, a characteristic less attributable to silent and immobile electronic monitoring devices. Often the presence and purpose of these devices is underscored by the experimenter in providing instructions, corrective feedback, etc. to subjects who are placed "alone" in experimental rooms in social facilitation studies. Although the experimenter is not physically present, it is quite clear to subjects that the experimenter is capable of directly monitoring the relevant channel of responding, whether it is visual, aural, or both



(Henchy and Glass, 1968; Cottrell, Wack, Sekerak and Rittle, 1968; Zajonc and Sales, 1966). The possible result of this monitoring, of course, is the creation of social facilitation effects in the "alone" condition, a situation which makes nearly impossible the accurate assessment of the weaker audience or observation conditions in these studies.

The most direct examination of the consequences of experimenter "monitoring" or "presence" is available from the social facilitation research reported by Henchy and Glass (1968). In order to test the "mere presence" notion and their opposing evaluation contention, these authors employed a pseudo-recognition task in order to measure social facilitation effects. Subjects were individually tested under one of four different audience conditions: (1) subject alone in the experimental room; (2) subject observed by two peers of slightly lesser status; (3) subject informed that he would be filmed (by a visible camera) and tape-recorded so that his performance could be analyzed at a later point by psychologists; and (4) subject closely observed by two psychologists who were introduced as such (actually these two observers were the same as in condition 2 but with different introductions and attire). The authors found support for their hypothesis that only conditions 3 (filmed and taped) and 4 (two psychologists) would demonstrate social facilitation effects due to the evaluation apprehension supposedly elicited by the expert audiences in those two conditions. Also reported was a difference between

conditions 1 (alone) and 2 (peer audience) which approached significance. Henchy and Glass suggest that this unexpected difference was the consequence of some degree of evaluation apprehension evidently present in condition 2 (peer audience), although they present no evidence to support this interpretation. In this particular study, subject's verbal responses were quite obviously monitored by means of an intercom connecting the experimental room with the experimenter's adjoining room. This "monitoring" may have enhanced performance in the "alone" condition, and thus decreased the differences between conditions 1 (alone) and 2 (peer audience).

More importantly, however, Henchy and Glass neglected two other possible conclusions with regard to their results:

(1) The strength of the "filmed and taped" condition strongly suggests that "alone" conditions, especially in social facilitation studies, have the potential to arouse evaluation apprehension, the strength of which is dependent upon the particular observational procedure employed. Most importantly, this study demonstrates that the labeling of experimental conditions as "subject alone" is an inadequate conceptualization; that the experimenter's ability and explicitness in monitoring subject behavior could be major determinants of the arousal of evaluation apprehension independent of his physical presence.

(2) The largest degree of social facilitation (and, apparently, evaluation apprehension) was brought about by the



presence of two observing psychologists; this was followed closely by the effect that was created when subjects were told that psychologists would later analyze the films and tapes being made of the experimental session. The significance of this study is quite clear; it appears that psychologists themselves may create a large degree of evaluation apprehension by their close observation of subjects in the experimental laboratory. This result lends further credence to Rosenberg's (1969) contentions that: (a) the experimenter is a very special type of observer, in that subjects attribute "special abilities . . . to those whose work is perceived as involving psychological interests and skills" (p. 281); (b) any aspect of the experimenter or of the experimental situation "that adds some further implication of interest in psychological evaluation will tend to increase the influence of the evaluation apprehension dynamic upon the subject's experimental responding" (p. 310).

It should be noted, however, that the effects of experimenter monitoring in any particular study are dependent to some degree upon the nature of the experimental task, also. If the task is a verbal one, for example, the subject may be concerned primarily with the degree to which the experimenter is able to hear his responses (aural monitoring). If, on the other hand, the task is one which can be immediately evaluated by visual inspection, then the subject might be more concerned with the experimenter's ability to see his behavior.

This position can be summarized by stating that task-relevant monitoring should be of more concern to the subject than task-irrelevant monitoring. Even if the experimenter is "present" with the subject, these variations are possible. Given that the task is one which might be monitored visually, the experimenter conceivably could be present in the same room, but at the same time prevented from viewing the subject due to a partition separating them. In this case, one would expect the subject to be less concerned about the "presence" of the experimenter than in a situation in which the experimenter was ostentatiously watching the subject's performance. Consistent with the Henchy and Glass (1968) position, it is assumed that subjects are concerned about the evaluation of their performances, and this condition leads to social facilitation effects. Experimentally, however, the logic is reversed: if varied degrees of experimenter monitoring are able to create differential social facilitation effects, then it is logically possible to infer the differential arousal of evaluation apprehension.

The present study, then, was designed to examine the above contentions by comparing the effects of varied degrees of experimenter monitoring (no monitoring, experimenter absent from the room; aural monitoring, experimenter present but behind a screen; visual monitoring, experimenter present and observing) upon performance at a simple laboratory task. It was predicted (hypothesis I) that increasing obtrusiveness of



experimenter monitoring, and the resulting evaluation apprehension, would produce successively higher performance scores at a vowel cancellation task (Allport, 1924). Due to the possibility that this effect might be moderated by the level of evaluation apprehension elicited by the instructions for the task, two sets of instructions were employed, one designed to elicit little or low apprehension, the other one calculated to arouse high apprehension. Consistent with these instruction sets, it was predicted (hypothesis II) that the high apprehension instruction set should produce higher performance scores than the low apprehension set. The possibilities for instructions X experimenter monitoring interactions are considerable. It was entirely possible that a ceiling effect might occur: subjects with the high apprehension instructions might be so concerned about their performance on the task that differing degrees of experimenter monitoring might have only a minimal influence upon performance. Conversely, a basement effect might occur with the low apprehension instructions: if subjects were not at all concerned about their performance, then it might not matter at all that the experimenter was observing them. Due to the ambiguity of the problem, no specific predictions were made with regard to possible instructions X experimenter monitoring interaction effects in this study. The existence of an interaction would provide primarily additional information as to the mediating effects of the instruction set; the absence of an interaction would



suggest that the evaluation apprehension elicited by the instruction set is additive to that aroused by the obtrusiveness of the experimenter's monitoring. Predictions were made only with regard to the performance scores at the vowel cancellation task, and thus no predictions were made regarding the number of errors, although it is generally accepted that the number of errors positively correlates with the rate of performance with this type of task (Dashiehl, 1935).

Lastly, it was predicted (hypothesis III) that the variance of scores on the vowel cancelling task would be smaller when subjects are visually monitored by the experimenter than in the other conditions. This constriction of the range of scores is most probably due to the fact that subjects in this situation are motivated to attend more closely to the task and thus tend to produce more uniform performance scores. Some peripheral evidence exists for this type of effect in the co-action research with judgements of weights and odors (Allport, 1924).

### Subjects

The subjects in this study were 96 undergraduate males solicited from courses which offered course credit for experimental participation. Subjects were drawn from both summer (N=48) and fall (N=48) subject pools.

### Design

Two independent variables, experimenter monitoring (three levels) and apprehension level of instructions (two levels) were combined in a 3 X 2 between subjects design (16 subjects/cell). The major dependent variable consisted of the total number of vowels canceled during a five minute interval. The total number of errors, consisting of vowels skipped or missed, was also recorded. In addition, subjects' responses regarding their reported level of apprehension were obtained from a final questionnaire and were also included in the analysis.

### Experimental task and materials

The vowel cancellation task employed in this experiment was previously used in a study of coaction (audience) effects, in which the "audience" condition resulted in an enhancement of group performance relative to an "alone" condition (Allport, 1924). Three legal sized sheets of paper with single-spaced typewriting were provided to each subject. The typewritten material consisted of letters assembled with no apparent order. No time limit was specified for this task, but subjects were instructed to work as fast as possible.

## Procedure

The small experimental room was plain and contained only two desks with chairs; a portable partition, when not in use, was placed against one of the side walls. When the subject entered the room (no mirrors, microphones or cameras), the experimenter explained: "In order to eliminate the possibility of the occurrence of certain biasing factors, all of the necessary instructions for this session are presented by means of videotape. Why don't you have a seat over there and we can begin." The subject was then seated at one of the desks in the room where he found the task materials placed face down. At this point, one of the three experimenter monitoring conditions was begun:

(1) No monitoring condition. (Experimenter absent during task performance)

In this condition the experimenter explained that he must leave to meet with another subject who had already begun the experimental task, but that he would return. The experimenter then turned on the videotape and left the room, closing the door behind him. After a little more than five minutes had passed (when the subject had finished with the experimental task), the experimenter returned to the room.

(2) Aural monitoring condition (Experimenter present in room but behind a screen; thus the experimenter could hear the subject but could not see him performing)

Under this condition, the experimenter explained that he was scoring some materials from previous subjects. Then the experimenter turned on the videotape, sat at the second desk behind the screen and began to make paperwork noises.

(3) Visual monitoring condition. (Experimenter closely observing subject's performance at the task)

Here the experimenter turned on the videotape and then sat on top of the second desk. The experimenter intensely observed the subject's performance and periodically took notes on a clipboard.



In addition to the three possible conditions of experimenter monitoring, the subject also experienced one of two videotapes providing instructions, one designed to elicit a high degree of apprehension regarding the task, the other, a low degree. The videotape scripts read as follows:

(1) High apprehension instructions

"Hello. The purpose of this experiment is to gather information about the way different individuals perform at a particular experimental task. Despite the simple appearance of this task, individual performance at this task has been shown by previous research to be related to a number of crucial personality variables, including intelligence. Of course, I can't tell you more about the task until you have completed it, so let's begin. Turn over the materials on the desk in front of you. The task is to cancel all the vowels (a, e, i, o, and u) on the pages in front of you, beginning with page one. In other words, draw a vertical line through each vowel that you come to. As an example, the first line of page one has been done for you. As you can see, there is no pattern to the letters on the pages. Work as rapidly as possible. Begin."

(2) Low apprehension instructions

"Hello. The purpose of this session is to gather normative information about the way that people perform at a particular experimental task. I'll be averaging the performances of a large number of people so that it will be possible to know the overall norms of performance for a large segment of the population. Later, this information will be used for cross-cultural comparisons of literate societies with differing lingual notation systems. Turn over the materials on the desk in front of you. (proceeds as above)

After the instructions were presented, the videotape screen was blank for exactly five minutes. At that point, the experimenter reappeared on the screen and announced: "Stop! Regardless of how much of the task you have finished, turn the test materials over and stop working." (In the no moni-

toring condition, the experimenter would then return to the room.) The experimenter then administered the following questionnaire and then debriefed the subject.

### Questionnaire items

The following items were included in the final questionnaire, in an attempt to provide self-report data regarding subject apprehension in the different conditions:

1. How concerned were you about your performance at the task?  
(put an "X" in the space which best indicates how you felt)  
not at all concerned :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: very concerned
2. How concerned were you about how well you were doing on the task?  
very concerned :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: not at all concerned
3. How concerned were you about the speed at which you worked at the task?  
not at all concerned :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: very concerned
4. How concerned were you about skipping or missing vowels?  
very concerned :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: not at all concerned
5. Did you find yourself going back to check for accuracy?  
very often :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: not at all
6. Did you feel that the experimenter was judging your performance?  
no, not at all :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: yes, very much



## Results

Due to the fact that both summer (N=48) and fall (N=48) subject pools were employed in this study, an initial analysis was performed to test for possible differences between these two populations. Analysis of variance failed to disclose any differences for the performance variable (total number of vowels cancelled), and thus the data for the two populations were pooled.

### Manipulation checks

In order to assess the effectiveness of the experimenter monitoring manipulation, subjects were asked in the final questionnaire "Where was the experimenter while you worked on the task?" All 96 subjects correctly answered this question. A second question, "What was the expressed purpose of this experiment?", was employed to examine the impact of the two instruction sets. Once again, all subjects could recall to at least some degree that the task was either related to personality and/or IQ, or represented part of a cross-cultural study.

### Performance scores

The performance score data were subjected to a three by two analysis of variance. The means and variances of the performance scores are presented in Table 1. Consistent with hypothesis I, there was a main effect for experimenter monitoring ( $p = .043$ ). Analysis of simple effects demonstrated, however, that this main effect was due solely to difference be-

Table 1

Means and Variances for the Number of Vowels  
Cancelled During A Five Minute Interval

Instructions  
Apprehension Level

		High	Low	
Experimenter Monitoring	No Monitoring	340.25 (4434.1) <sup>a</sup>	361.00 (3823.1)	350.63 (4122.9)
	Aural Monitoring	378.75 (5178.5)	389.56 (3776.7)	384.16 (4450.0)
	Visual Monitoring	392.38 (3854.0)	373.81 (3088.7)	383.09 (3562.3)
		370.46 (3854.0)	374.79 (3088.7)	373.63 (3562.3)

<sup>a</sup> = variance

tween the mean for the no monitoring condition (350.63) and the pooled mean of the other two monitoring conditions (383.63). Thus, the combined mean number of vowels cancelled for those conditions in which the experimenter remained in the room was significantly higher than the mean for the condition in which the experimenter was absent from the room.

Contrary to the prediction of hypothesis II, it is clear from Table 1 that no differences in performance scores occurred for the instructions variable. The mean for the high apprehension instructions (370.46) was not significantly different from the mean for the low apprehension instructions (374.79).

#### Error Scores

The vowel cancellation task was so easily performed that errors were uncommonly few. In fact, less than one-half of all subjects had any errors at all. Error scores were found not to correlate with performance scores or questionnaire items and thus were not included in the statistical analysis presented here.

#### Variances

Examination of the variances in Table 1 discloses that within the visual monitoring condition the cell variances (2274.3, 1887.2) are much smaller than those in the remaining four cells (5178.5, 3776.7, 4434.1, 3823.1). In order to test hypothesis III, that there should be less performance score variance in the visual condition than in other condi-



tions, the average variance in the first two cells (2076.3) was compared with the average variance of the latter four cells (4286.5) in a simple F ratio. This test demonstrated that the average variance in the visual monitoring condition was significantly smaller ( $p < .05$ ) than the average variance of the remaining two conditions, thus confirming hypothesis III.<sup>2</sup>

In both of the individual instruction conditions, this difference in variance for observed versus unobserved subjects also approached significance ( $p < .10$ ). To simplify the presentation of these differences, the relevant frequency polygons are presented in Figure 1. In order to provide an adequate comparison of the variances of the three experimenter monitoring conditions, all of the scores in the no monitoring condition were increased by 33.00. The adjusted mean performance score for this group then became 383.63, which is the average of the means for the other two conditions. As can be seen from this Figure, the smaller variances in the visual monitoring condition are due to contraction on both ends of the frequency distribution.

#### Questionnaire data

Analysis of subject responses to the six questionnaire items revealed that none of these items correlated significantly with performance scores. Despite this finding, analysis of variance disclosed two of the six items did show significant differences with regard to instructions variable ("How concerned were you about your performance on the task?",

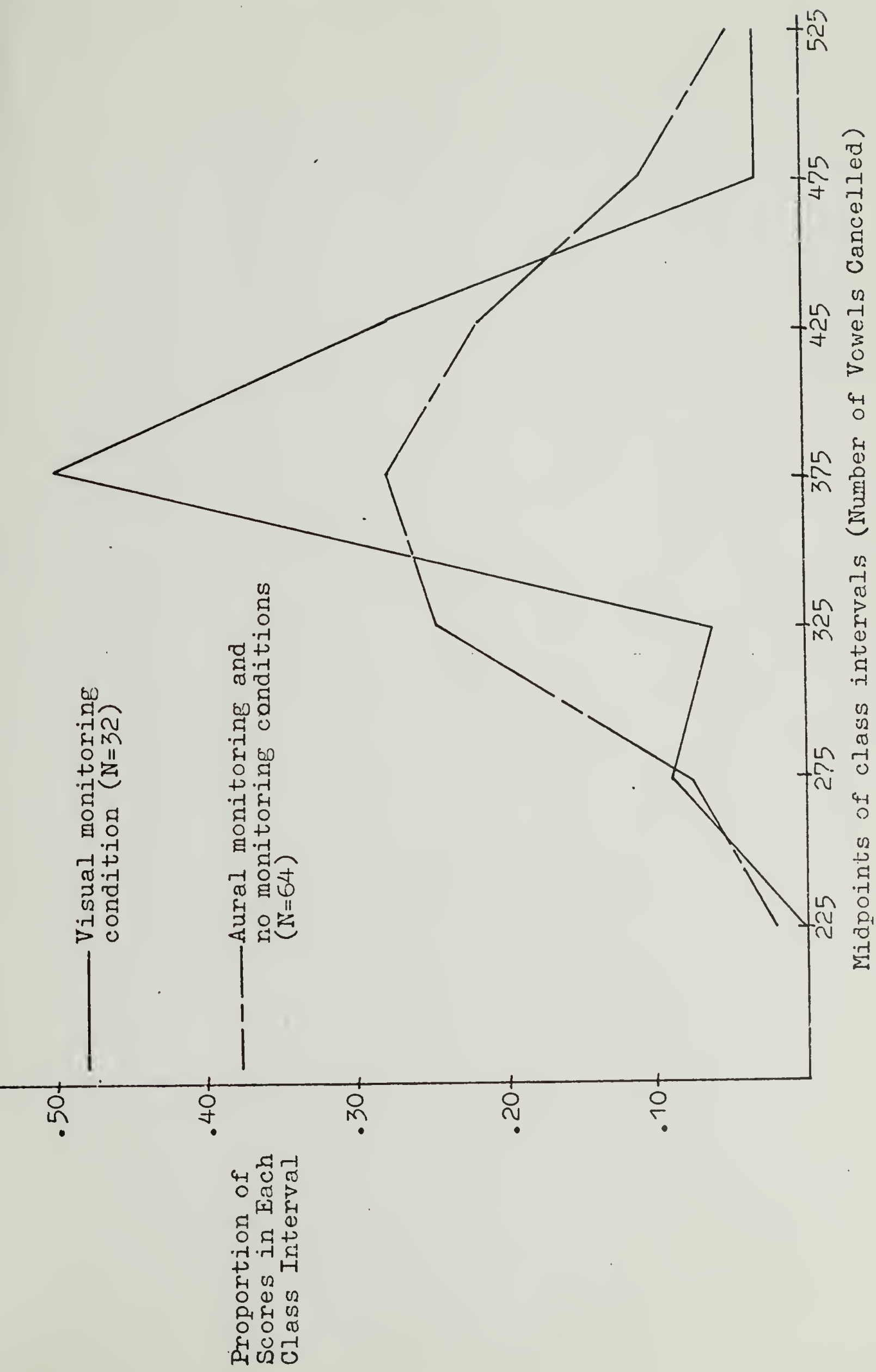


Figure 1. Frequency polygon for distribution of scores in the visual monitoring condition versus aural monitoring conditions (with adjusted no monitoring scores).

$p < .012$ ; and "How concerned were you about skipping or missing vowels?",  $p < .046$ ). A third item ("How concerned were you about how well you were doing on the task?" was marginally significant on this variable ( $p < .084$ ). As might be expected with these items, subjects were more concerned in the high apprehension instructions condition than in the low apprehension instructions condition.



## Discussion

Regardless of whether the experimenter observed the subject directly or sat behind a screen, the presence of the experimenter in the experimental room with the subject clearly resulted in significantly higher performance scores than when the experimenter was absent from the room. This result was taken as a disconfirmation of hypothesis I: that increasing obtrusiveness of experimenter monitoring should produce successively higher performance scores on a vowel cancellation task. Originally, it had been expected that the mean performance score in the aural monitoring condition would be less than the mean of the visual monitoring condition, but still larger than that of the no monitoring condition. In terms of the mean performance scores for these three conditions (Table 1), however, this is clearly not the case: when compared to the removal of the experimenter from the experimental room, the placement of the experimenter behind a screen increases mean performance as much as does direct observation by the experimenter.

It should be noted at this point that the high apprehension instructions condition is more typical of the experimental setting of social facilitation research than is the low apprehension instruction condition. In social facilitation studies, subjects are confronted with a situation in which they find themselves performing individually at a learning task. Given this situation, one would expect subjects to be

concerned that their performance on the task might reflect something about their intellectual or cognitive capabilities, as with the high apprehension instructions in the present study. Examining the mean performance scores for only the high apprehension instruction condition, one finds some suggestion that aural monitoring resulted in an intermediate level of performance (378.75), but tests of simple effects revealed that this mean was not significantly different from either the no monitoring (340.25) or visual monitoring (392.38) means.

In terms of the "experimenter monitoring" approach discussed earlier, the similarity of performance scores in both the aural and visual monitoring conditions could be interpreted to mean that these two conditions are nearly equivalent in the amount of evaluation apprehension they can elicit from subjects. Due to the fact that subjects performing the vowel cancellation task do make some noise in the otherwise silent experimental room, it is possible that subjects did in fact feel that their task performance was monitored, even when the experimenter sat behind a screen. Alternatively, it is conceivable that an experimenter sitting behind a screen represents just as threatening an evaluator as one directly observing the subject, in that the former is perhaps a more formal testing situation. In any event, regardless of underlying causes, it is clear that the physical presence of the experimenter in the same room as the subject increases perfor-

mance scores at a simple task, thus demonstrating social facilitation effects in the laboratory as a consequence of the experimenter's "presence."

This finding has direct relevance for social facilitation research, which has been continually characterized by hazy definitions of "subject alone" conditions. Generally, the potency of a given "audience" condition is ascertained by determining if scores in that condition are significantly different from scores in the "alone" condition. In this case, it should be evident that how one defines the "alone" condition directly determines the relative strength of a given "audience" condition. Given this state of affairs, it is not surprising that inconsistent and contradictory results are often found within the social facilitation literature, especially with the weaker "audience" conditions (e.g., "mere presence"). In attempting to rectify this situation, researchers in this area might standardize their definition of "subject alone" in terms of the observational procedures employed by the experimenter in their studies. The difficulty with this approach, however, as alluded to earlier, is that the definition of "subject alone" undoubtedly varies somewhat, depending upon the nature of the behavior required by the experimental task. For some tasks, visual monitoring by the experimenter is sufficient, while for others, aural monitoring may be required, or even both types of monitoring together. Alternatively, a more practical solution to this dilemma might be



to encourage researchers to report much more elaborate descriptions than presently exist of both their experimental tasks and the specific observational procedure employed. At the very least, this approach would allow more accurate comparisons of these studies, and would alert future researchers in this area to the possible problems of experimenter monitoring.

Hypothesis II was not confirmed. It had been predicted that the high apprehension instructions would produce higher performance scores than the low apprehension instructions. Examination of Table 1 shows no difference in the mean performance scores for the instructions manipulation (370.46, 374.79). This absence of performance differences occurs despite the fact that three of the six questionnaire items gave some indication that subjects self-reported apprehension levels were higher in the high apprehension condition. Given this state of affairs, it appears that a "basement effect" may have occurred here. In other words, both of the independent variables (experimenter monitoring, apprehension level of instructions) represent manipulations of evaluation apprehension, at least at the theoretical level. In this situation, it is entirely possible for the stronger variable to mask the effects of the weaker one. In a previously mentioned study by Criddle (1971b), this type of effect was also evident. Criddle found no significant differences between three varied "audience" conditions. His laboratory procedure,

however, included the physical presence of the experimenter across all conditions. The presence of the experimenter can be viewed as an additional "audience" manipulation, which was so strong as to "wash out" any differences between his actual "audience" conditions. That this type of confounding was the case is verified by other research by Criddle (1971a) in which significantly different "audience" effects were obtained using almost identical conditions. In this other study, though, the experimenter was absent from the experimental room in all conditions. Viewing these two studies together, one can gauge the strength of the experimenter's presence in terms of an "audience condition." Thus, the increased background level of evaluation apprehension resulting from the physical inclusion of the experimenter is sufficiently high enough to prevent the detection of effects due to weaker "audience" manipulations. In this way, a "basement effect" was created.

In the present study, a similar situation may have occurred. Since the instruction manipulation did not affect performance scores, but did affect indicators of subject apprehensiveness, it appears likely that the background level of evaluation apprehension (resulting from the subjects being in this type of experiment, performing at this particular task, etc.) was high enough as to eliminate any performance effects due to variations in the instructions set. Thus the subjects may have been so apprehensive about experimental performance

at a vowel cancellation task that any effects due to the receipt of additional information about the purpose of the task were "washed out." In other words, the instruction manipulations were not effective because the background level of apprehension pre-existing within the experimental setting was so high that any differences due to varied instructions were undetectable.

Lastly, it was found that groups of subjects who were directly observed by the experimenter produced significantly lower performance score variance than did unobserved subjects. This result confirms hypothesis III, that the variance of performance scores would be smallest when subjects were visually monitored by the experimenter, although with some reservations. The strength of this result is diminished somewhat by the fact that the necessary comparison required the pooling of several conditions, thus generating large values of N. On the other hand, within each of the instruction conditions, the differences in variances for observed versus unobserved subjects reached marginal significance ( $p < .10$ ). In any event, this result is perhaps the most interesting aspect of this study: Figure 1 clearly shows how both ends of the distribution contract for observed subjects. The original explanation for this effect, discussed earlier, suggested that the direct visual monitoring of the experimenter would motivate subjects to attend more carefully to the experimental task. This rationale conveniently explains why the low end



of the distribution is contracted with direct observation, but is less satisfactory in attempting to understand why the high end of the distribution also contracts. More specifically, attending to the task should increase the performance scores of subjects who might otherwise be distracted by other aspects of the experimental situation; on the other hand, it is more difficult to argue that attending to the task, for other subjects, should decrease their performance scores, especially with such an easy task. A more productive explanation might be the following: subjects working unobserved at a simple laboratory task vary in the degree to which they are affected or respond to evaluation apprehension. Some subjects are overly concerned about the possible consequences of their performance and work intensely at the task with an extreme degree of involvement, while other subjects only minimally care about how they perform in some silly experiment, but are still willing to comply with the basic groundrules of the experiment in order to obtain their experimental credit slip. As a consequence, this group produces performance scores with a great deal of variation. If a similarly composed group of subjects now work at the task while closely observed by the experimenter, their degree of concern with their task performance must now be tempered by evaluation apprehension about acting appropriately in the experimental setting; subjects are now concerned with the attributions made about them by the experimenter on the basis of their immediate social behavior.

Although the experimenter will eventually obtain a measure of their task performance in either situation, only when the experimenter is present with the subject is the way in which the subject performs the task open to scrutiny. Given this situation, subjects who would otherwise be lax in their performance are more prone to work "as fast as possible"; normally over-zealous subjects are reluctant to reveal too much personal involvement in the experimental task and thus are inhibited from racing along as if competing. In this way the distribution might be significantly narrowed.

Regardless of the specific mechanism by which performance score variance is reduced by a closely observing experimenter, the effect itself suggests a number of methodological implications. In terms of social facilitation research, it is evident that the inclusion of an "alone" condition which involved no experimenter monitoring could increase substantially the size of the overall error term by which the effects of different "audience" conditions are evaluated. The result of increased error variance upon significance testing is unclear in this particular situation, since the use of this "alone" condition also increases the relative size of the social facilitation effects in the various "audience" conditions. A related problem arises in regard to the present study. If only one (instead of two) conditions of non-observation had been included in the present design, the overall error term would have been diminished considerably, thus increasing the

the probability of obtaining statistically significant results. Thus the elimination of the "aural monitoring" condition (with its high performance score variance) from statistical analysis, for example, results in a much stronger main effect for the experimenter monitoring. In this type of situation, then, a researcher may find himself torn between examining more than one condition of non-observation and attempting to reduce error variance in the overall experiment.

On a more general level, the occurrence of greater variances under conditions of non-observation may be related to a larger methodological problem within psychology: the interaction between observer and observed. In this study, greater variances arose when subjects performed unobserved at a simple laboratory task. Milgram (1965) has previously concluded that "obedience dropped sharply as the experimenter was physically removed from the laboratory." And Allport (1924) long ago demonstrated that groups of subjects show more variance in their judgements when alone than when in the presence of coactors. Although the evidence is neither voluminous nor conclusive, the available data suggest that one effect of being observed or monitored by others may be the constriction of variation in the particular social behavior under scrutiny. Allport (1924) referred to this tendency in the individual as the "attitude of social conformity," which he defined as "the often unconscious basic human tendency to temper one's opinions and conduct by deference to



the opinions and conduct of others (p. 278)." Obviously, there is no reasonable argument for the exclusion of experimenters from the category of "others." Despite this, however, little is known in psychology about the effects of experimenter monitoring upon the performance of subjects. Even within the social facilitation area, this area has remained relatively neglected.

It is the present opinion of this author that further research in experimenter monitoring will demonstrate that much of the "controlled situation" generally attributed to the laboratory setting is, in fact, partially the consequence of the monitoring of subjects by experimenters, especially when the dependent variable represents some form of social behavior. One of the explicit goals of conducting research within the laboratory is the reduction of unwanted error variance. Although the generalizability of the present study is difficult to ascertain, its results, when taken with other evidence, suggest that the observing presence of the experimenter within the laboratory may contribute to this goal. In this light, it may not be surprising that field experiments, in which subjects have no awareness of being monitored by experimenters, often replicate laboratory results only with great difficulty and borderline levels of significance (due to larger error terms).

Such speculation is no substitute for research in this area, however. Before the epistemological status of the

laboratory experiment can be thoroughly and convincingly discussed, more information is needed about the effects of observation in general, and experimenter monitoring in particular. The variance phenomenon should be examined with a broad range of experimental tasks and experimental paradigms. Although the present study represents one method of assessing experimenter monitoring effects, a much more satisfactory design would include a condition in which the subject performs a certain task with no knowledge of being in an experiment (and thus unaware of being monitored by an experimenter). The difficulties here are at least threefold: (1) identifying a behavior which could be used as an index of social facilitation, but which could be observed in naturalistic settings (the pseudo-recognition task is clearly insufficient here); (2) developing a definition, both conceptually and practically, of what is meant by "subject alone"; and (3) attempting to estimate the absolute value of the effect of experimenter monitoring (this is an illusory goal--this value will always be relative to the naturalistic "alone" condition chosen by the researcher). Nevertheless, it is important to point out that any valid attempt to assess the effects of experimenter monitoring upon subject performance demands stepping out of the laboratory situation, at least in terms of the social phenomenology of the subject. This journey is required because "being in an experiment" necessarily implies to the subject that

he or she is being monitored by an experimenter, at least to some extent. Thus leaving the laboratory (however temporarily) becomes one prerequisite for understanding what is occurring within the laboratory.



### Summary

The present study was designed to assess the effects of varied degrees of "experimenter monitoring" and two different instructions sets upon subject performance at a simple vowel cancelling task. Three conditions of "experimenter monitoring" were developed from a social facilitation framework (no monitoring, experimenter absent from the room; aural monitoring, experimenter sitting behind a screen in the same room; visual monitoring, experimenter observes directly). Two videotaped instruction sets were employed, one designed to elicit a high degree of apprehension about the task, the other designed to create a low degree. These two variables were combined in a 3 X 2 between subjects design. In all, 96 subjects individually participated in the study (16 subjects/cell). The major dependent variable consisted of the total number of vowels cancelled during a five minute interval. In addition, subjects' responses regarding their reported levels of apprehension were obtained from a final questionnaire, the results of which were included in the analysis.

Hypothesis I stated that increased experimenter monitoring would result in increased performance scores. This hypothesis was only partially supported and was considered disconfirmed. Although a significant main effect occurred for experimenter monitoring, both the aural and visual conditions resulted in nearly identical increases in performance scores when compared with the no monitoring condition. Thus social

facilitation effects as a consequence of the experimenter's "presence" were demonstrated, but this finding did not support the "experimenter monitoring" approach.

Hypothesis II predicted higher performance scores for the high apprehension instructions than for the low apprehension instructions. No differences in performance scores were found with this variable. It was argued that the lack of differences for this variable were due to a "basement effect" inherent in the experimental procedure.

Hypothesis III maintained that groups of subjects who were visually monitored by the experimenter would show significantly less performance score variance than subjects who were either aurally monitored or not monitored at all. This hypothesis was confirmed and the resulting methodological implications were discussed.

## Footnotes

<sup>1</sup>The term distal is used here to denote Criddle's concentration upon aspects within the environment as causative factors, as opposed to the proximal causes generally invoked to "explain" social facilitation (e.g., the motivational state of the subject).

<sup>2</sup>Although a significant difference between variances was found in this case, it does not necessarily represent a violation of the homogeneity of variance assumption required by analysis of variance. Hartley's  $F_{\max}$  test (Winer, 1962, p. 206) was applied to the data and was found to be nonsignificant, thus indicating that no major violation of the homogeneity assumption had occurred. In addition, Myers (1972) and others indicate that the F distribution remains very robust with heterogeneity of variance, provided that (1) scores are approximately normally distributed, and (2) that equal ns exist. The effect of heterogeneity of variance under these conditions is a mild inflation of alpha levels. Myers also cites Box (1954), who demonstrated that under the conditions stated, that a 20:1 variance ratio was necessary to inflate a .05 level to .07. In the present study, the greatest ratio for any two cells was 2.75, with the main effect significant at the .043 level. Given this situation, it seems that heterogeneity of variance would probably not inflate the alpha level above .05.



## Bibliography

- Allport, F. H. Behavior and experiment in social psychology. Journal of Abnormal and Social Psychology, 1919, 14, 297-308.
- Allport, F. H. Social psychology. Boston: Houghton Mifflin Company, 1924.
- Bergum, B. O. and Lehr, D. J. Effects of authoritarianism on vigilance performance. Journal of Applied Psychology, 1963, 47, 75-77.
- Box, G. E. P. Some theories on quadratic forms applied in the study of analysis of variance problems: I. Effect of inequality of variance in the one-way classification. Annals of Mathematical Statistics, 1954, 25, 290-302.
- Cottrell, N. B., Rittle, R. H., and Wack, D. L. The presence of an audience and list time as joint determinants of performance in paired associate learning. Journal of Personality, 1967, 35, 425-434.
- Cottrell, N. B., Wack, D. L., Sekerak, G. J., and Rittle, R. H. Social facilitation of dominant responses by the presence of an audience and the mere presence of others. Journal of Personality and Social Psychology, 1968, 9, 245-250.
- Cottrell, N. B. Performances in the presence of other human beings: Mere presence, audience and affiliation effects. In E. C. Simmel, R. A. Hoppe and G. A. Milton (Eds.), Social facilitation and imitative behavior. Boston: Allyn and Bacon, Inc., 1968.

- Criddle, W. D. The physical presence of other individuals as a factor in social facilitation. Psychonomic Science, 1971, 22, 229-230. (a)
- Criddle, W. D. Determinants of social facilitation in humans. Dissertation Abstracts International, 1971, 31B, 7569-7570. (b)
- Criddle, W. D. One-way screen observation of psychiatric patients. Psychological Reports, 1971, 28, 918. (c).
- Dashiell, J. F. Experimental studies of the influence of social situations on the behavior of individual human adults. Handbook of Social Psychology, 1935, 2, 1097-1158.
- Ekdahl, A. G. Effect of attitude on free word association time. Genetic Psychology Monographs, 1929, 5, 253-338.
- Fraser, D. C. The relation of an environmental variable to performance in a prolonged visual task. Quarterly Journal of Experimental Psychology, 1953, 5, 31-32.
- Ganzer, F. Effects of audience presence and test anxiety in learning and retention in a serial learning situation. Journal of Personality and Social Psychology, 1968, 8, 194-199.
- Giogi, A. Psychology as a human science--a phenomenologically based approach. New York: Harper and Row, 1970.
- Henchy, T. and Glass, D. C. Evaluation apprehension and the social facilitation of dominant and subordinate responses. Journal of Personality and Social Psychology, 1968, 10, 446-454.

Ichheiser, G. Ueber die Veranderung der Leistungsbereitschaft durch das Bewusstsein einen Zuschauer zu haben. Psychotechnische Zeitschriften, 1930, 5, 52-53.

Leventhal, H. and Fischer, K. What reinforces in a social reinforcement situation--words or expressions? Journal of Personality and Social Psychology, 1970, 14, 83-94.

Martens, R. Effect of an audience on learning and performance of a complex motor skill. Journal of Personality and Social Psychology, 1969, 12, 252-260.

Meddock, T. D., Parsons, J. L. and Hill, K. T. Effects of an adult's presence and praise on young childrens' performance. Journal of Experimental Child Psychology, 1971, 12, 197-211.

Milgram, S. Some conditions of obedience and disobedience to authority. Human Relations, 1965, 57-76.

Myers, J. L. Fundamentals of Experimental Design (second edition) Boston: Allyn and Bacon, Inc., 1972, 72-73.

Pessin, J. The comparative effects of social and mechanical stimulation on memorizing. American Journal of Psychology, 1933, 45, 263-270.

Peterson, R. F. and Whitehurst, G. J. A variable influencing the performance of non-reinforced imitative behaviors. Journal of Applied Behavioral Analysis, 1970, 4, 1-9.

Rosenberg, M. J. The conditions and consequences of evaluation apprehension. In R. Rosenthal and R. Rosnow (Eds.), Artifact in behavioral research. New York: Academic Press, 1969, 279-349.



- Shrauger, J. S. Self-esteem and reactions to being observed by others. Journal of Personality and Social Psychology, 1972, 23, 192-200.
- Wapner, S. and Alper, T. The effect of an audience on behavior in a choice situation. Journal of Abnormal and Social Psychology, 1952, 47, 222-229.
- Zajonc, R. Social facilitation. Science, 1965, 149, 269-274.
- Zajonc, R. and Sales, S. Social facilitation of dominant and subordinate responses. Journal of Experimental Social Psychology, 1966, 2, 160-168.



